Genomic sequence of HIV-1 from four members of the Sydney Blood Bank Cohort of Long-term Non-progressors

Sequence Note

Robert Oelrichs[1*], Anna Tsykin[1], David Rhodes[1], Ajantha Solomon[1], Anne Ellett[2], Dale McPhee[2] and Nicholas Deacon[1].

AIDS Molecular Biology Unit[1] and AIDS Cellular Biology Unit[2], Macfarlane Burnet Centre for Medical Research, Yarra Bend Road, Fairfield, 3078, AUSTRALIA.

* To Whom correspondence should be addressed.

Running Title: Sydney Blood Bank Cohort HIV-1 Genomic Sequences

The Sydney Blood Bank Cohort of long term non-progressors (SBBC) was first reported in 1992 as a collection of six people who were infected with HIV-1 between 1982 and mid-1984 via blood products from a common donor (1). Long-term follow up of five of the six recipients and the donor had shown persistent asymptomatic HIV-1 infection with no signs of progression to disease. The sixth recipient, C83, had died of Pneumocystis carinii pneumonia thought to be the result of immunosuppressive therapy for systemic lupus erythematosis. Further tracing has since identified two more recipients (C124 and C135) who are now included as members of the SBBC (2). These recipients were also asymptomatic and showed no signs of HIV/AIDS progression. Of the eight HIV-1 positive recipients identified, C124 and C18 have also died from causes unrelated to HIV/AIDS.

HIV-1 sequences from either cultured virus isolates or patient peripheral blood mononuclear cells had similar deletions in the *nef* gene, in the region of overlap of *nef* and the U3 region of the long terminal repeat (LTR) (3). Full-length sequencing of one isolate genome and amplification of selected HIV-1 genome regions from other cohort members revealed no other abnormalities of obvious functional significance. The strain of HIV-1 shared by the SBBC thus represents an attenuated, non-pathogenic form of the virus and has potential for use as a live HIV-1 vaccine. This hypothesis is strengthened by the similarly attenuated nature of forms of the simian immunodeficiency virus (SIV) that have been engineered with deletions in the *nef* gene (4,5). Infection of macaque monkeys with these attenuated viruses fails to cause the usual progression to AIDS and protects the animals from subsequent challenge with a pathogenic strain of SIV.

In order to more thoroughly confirm that the attenuated nature of the SBBC HIV-1 was due only to the characterised *nef*/LTR deletions, full length sequence of the HIV-1 genome was determined from four cohort members: C18 (from two independent sources distinct from the originally reported sequence - an infectious molecular clone and directly from a biologically cloned isolate); D36 (the cohort donor); C54 and C98. In each case the source of virus was a Hirt supernatant (6) from a low-passage number culture of PHA-activated donor peripheral blood mononuclear cells (PBMC); excepting C54, where the source used was a lysate of patient PBMC, prepared as described (3).

Full-length HIV-1 DNA sequence was determined by near full-length PCR (9.11 kb) between base positions corresponding to pNL4-3 621 and 9731, followed by amplification of a series of 14 overlapping PCR products. The oligonucleotide sequences of these primers were based on consensus sequences from conserved regions of HIV-1 subtype B and tailed with M13 forward and reverse universal primer sequences to enable direct automated sequencing of the PCR products from both strands. This rapid technique for sequencing full-length HIV-1 subtype B is detailed elsewhere (7). SBBC HIV-1 sequences have been deposited in the Genbank data base, Accession numbers C18 biological clone: AF042102, pC18RO1: AF042106, D36: AF042105, C54: AF042103, C98: AF042104.

A phylogenetic tree was derived from the above five SBBC HIV-1 genomic sequences and all full-length sequences of subtype B HIV-1 available from the HIV sequence

database as updated on May 28th, 1997 (8), Figure 1. The sequences are approximately 8 kb in length, extending from the UTR 5' of *gag* to the end of *env* (nucleotides 668-8782 in pNL4-3). The *nef*/LTR has been excluded from analysis as extensive deletions and rearrangements in the cohort viruses make unbiased alignment difficult. All computations were done through the Australian National Genomic Information Service (ANGIS), WAG interface. The alignment was performed using the CLUSTAL algorithm and the tree shown was constructed using Distances (with Jukes-Cantor corrections) and Growtree (neighbour joining method), GCG package. The use of maximum likelihood algorithm DNAML and parsimony method DNAPARS, both ported from the PHYLIP package, resulted in trees with no significant differences. Bootstrap values shown (for 100 replicates) were generated using SEQBOOT, DNADIST, NEIGHBOR and CONSENCE, also ported from the PHYLIP package.

SBBC viruses group with maximum certainty and are positioned closely with MBC200, which is the only other full-length Australian HIV-1 sequence available (7), as well as with MN, LAI, LW123, HXB2 and TH475. Phylogenetic analysis performed on smaller regions shows that the SBBC viruses group within subtype B in p17, *pol* and *env*, providing no evidence of viral mosaicism at a gross level (Figure 2).

A consensus using the described 8 kb data was generated and compared to the consensus subtype B sequence (8). Nucleotides entirely conserved across subtype B that were uniformly mutated in SBBC viruses are listed in Table 1 and changes in the predicted gene product noted where the mutations are non-synonymous.

None of the consistent nucleotide mutations corresponds to known splice donor or acceptor sites (9,10,11). However, it is worth noting three splice site mutations that are conserved in the three viral isolates from C18. The major splice donor at pNL4-3 base position 4962 is disrupted by the mutation T to C at position 4964, the only example of this in HIV-1 subtype B. There is inactivation of two splice acceptor sites by mutations at positions 6602 (A to T) and 6610 (G to T). Mutation at the latter site in pNL4-3 has resulted in a virus that produced wild-type quantities of spliced messenger RNA and protein but was non-infectious in PBMC culture (10). The conservation of these point mutations across three independant isolates from this patient indicates evolutionary pressure, perhaps exerted by a toxic or dominant-negative effect of the truncated Nef protein in this virus.

None of the predicted amino acid changes has been previously described as being linked to LTNP. The conservation across four SBBC member viruses of the additional three amino acids at the C terminus of the predicted Rev protein may also indicate a selective response to the different evolutionary pressures acting on *nef*-deficient HIV-1 *in vivo*; however these data confirm the previous observation that no significant differences link the viruses of the SBBC, or account for their attenuated phenotype, other than the deletions and rearrangements of the *nef*/LTR region (3).

Acknowledgments

The authors thank Dr. Francine McCutchan for her helpful discussions concerning these data, Dr.s Carolyn Bucholtz and Timothy Littlejohn for assistance with the phylogenetic analysis and Dr. Nelson Michael for advice concerning long-range PCR.

References

1. Learmont J, Tindall B, Evans L, Cunningham A, Cunningham P, Wells J, Penny R, Kaldor J and Cooper DA. Long-term symptomless HIV-1 infection in recipients of blood products from a single donor. Lancet 1992;340:863-67.

2. Learmont J, Cook L, Dunckley H and Sullivan J: Update on long-term symptomless HIV-1 infection in recipients of blood products from a single donor. AIDS Res Hum Retroviruses 1995;11(1):1.

3. Deacon N, Tsykin A, Solomon A, Smith K, Ludford-Menting M, Hooker D, McPhee D, Greenway A, Ellett A, Chatfield C, Lawson V, Crowe S, Maerz A, Sonza S, Learmont J, Sullivan J, Cunningham A, Dwyer D, Dowton D and Mills J. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science 1995;270(5238):988-91.

4. Daniel M, Kirchhoff F, Czajak S, Sehgal P and Desrosiers R. Protective effects of a live attenuated SIV vaccine with a deletion in the *nef* gene. Science 1992;258(5090):1938-1941.

5. Igarashi T, Ami Y, Yamamoto H, Shibata R, Kuwata T, Mukai R, Shinohara K, Komatsu T, Adachi A and Hayami M. Protection of monkeys vaccinated with *vpr*- and/or *nef*-defective simian immunodeficiency virus strain mac/human immunodeficiency virus type 1 chimeric viruses: a potential candidate live-attenuated human AIDS vaccine. J Gen Virol 1997 78(Pt 5):985-989.

6. Hirt, B. Selective extraction of polyoma DNA from infected mouse cell cultures. J Mol Biol 1967;36:365-369.

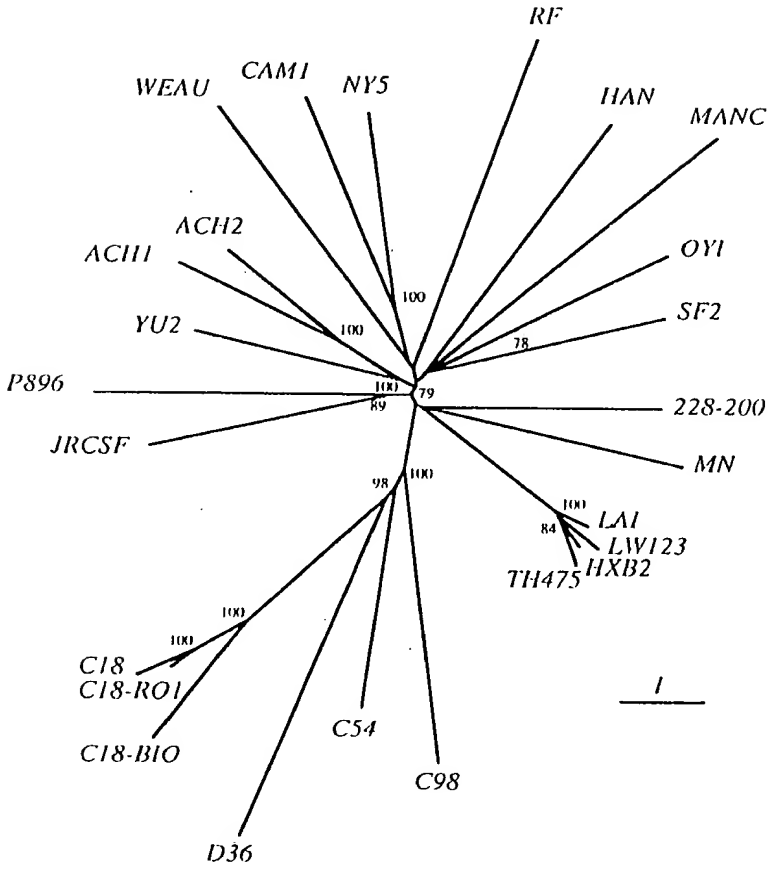
7. Oelrichs R, Coates Fryer K, Deacon N, Land S, Chatfield C and McPhee D. Biological characterisation and full-length genomic sequence of two Australian primary isolates of HIV-1. 1997. Manuscript in preparation.

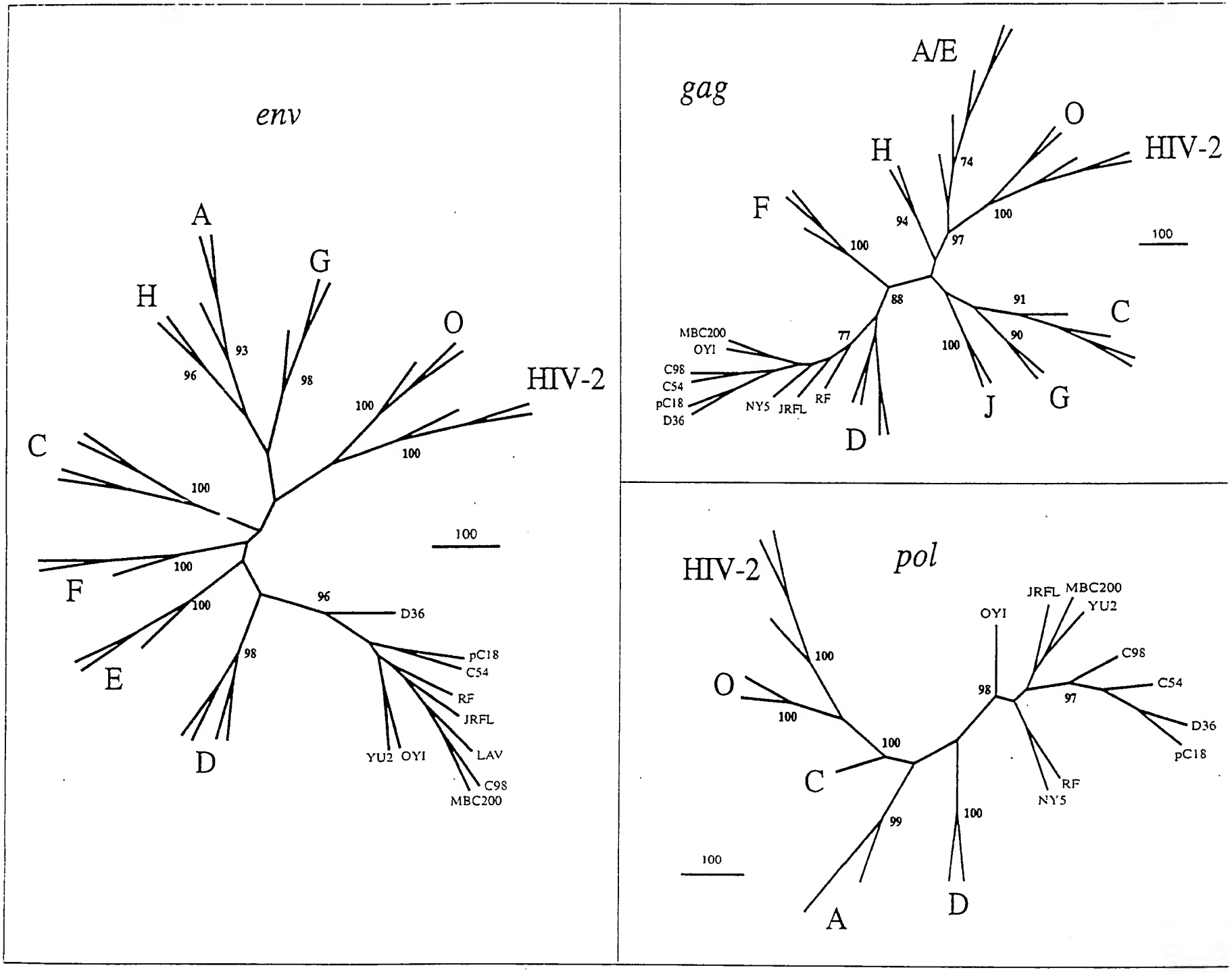
8. Myers G, Hahan B, Mellors J, Henderson L, Korber B, Jeang K, McCutchan F and Pavlakis G, Editors. Human Retroviruses and AIDS Database Compendium. Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, U.S.A. http://hiv-web.lanl.gov.

9. Schwartz S, Felber B, Fenyo E and Pavlakis G. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. J Virol 1990;64(11):5448-5456.

10. Purcell D and Martin M. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J Virol 1993;67(11):6365-6378.

11. Smith J, Azad A and Deacon N. Identification of two novel human immunodeficiency virus type 1 splice acceptor sites in infected T cell lines. J Gen Virol 1992;73(Pt 7):1825-1828.


RF
WEAU
CAM1
NY5
HAN
MANC
ACH2
ACH1
OYI
SF2
YU2
P896
228-200
JRCSF
MN
LAI
LW123
HXB2
TH475
C18
C18-RO1
C18-BIO
C54
C98
D36
100
100
78
100
89
79
98
100
100
84
100
100
1


env
A
G
H
O
HIV-2
C
F
E
D
93
96
98
100
100
100
100
100
100
96
98
100
D36
pC18
C54
RF
JRFL
LAV
C98
MBC200
YU2
OYI
gag
A/E
H
O
HIV-2
F
74
94
100
97
100
88
91
C
MBC200
OYI
77
100
90
C98
C54
pC18
D36
NY5
JRFL
RF
D
J
G
pol
HIV-2
JRFL
MBC200
YU2
OYI
100
C98
O
98
C54
97
100
D36
100
pC18
C
RF
NY5
99
100
A
D

Equivalent base position in pNL4-3	Mutation in SBBC viruses	Predicted amino acid mutation ("-" indicates no change)
1930	G → A	*gag* G381S
2426	G → A	-
2957	C → T	-
3284	G → A	-
3368	G → A	-
3813	T → C	-
4097	G → A	-
4121	G → A	-
4253	C → A	-
4670	T → C	-
5340	T → C	-
5732	A → G	-
5943	C → T	-
6183	A → G	-
6357	A → G	gp120 K18R
8053	T → C	-
8276	A → G	gp120 I658V
8641	T → G	gp120 N779K *rev* STOP → ECC
8698	G → A	-

Table 1: SBBC consensus sequence compared to subtype B consensus sequence. Nucleotides conserved across subtype B that are mutated in C18, C54, C98 and D36 are listed and, where non-synonymous, the predicted change in amino acid sequence is listed. Mutation T8641G results in loss of the *rev* termination signal and an additional three amino acids are predicted for the Rev protein from these viruses.

Figure 1: Genetic relationships of the SBBC viruses. An unrooted phylogenetic tree was developed using near full-length proviral sequences (minus the *nef*/LTR region) from SBBC viruses and available full-length sequences of subtype B HIV-1. A distance scale is shown, representing estimated number of substitutions per 100 bases. Significant bootstrap values are shown beside branchpoints (values out of 100). Sequence name abbreviations are as listed in the HIV Sequence database. SBBC sequences: C18 (previously reported sequence), C18-RO1 (molecular clone derived from C18), C18-BIO (biological clone derived from C18), D36, C54 and C98.

Table 1: SBBC consensus sequence compared to subtype B consensus sequence. Nucleotides conserved across subtype B that are mutated in C18, C54, C98 and D36 are listed and, where non-synonymous, the predicted change in amino acid sequence is listed. Mutation T8641G results in loss of the *rev* termination signal and an additional three amino acids are predicted for the Rev protein from these viruses.

Reprint requests should be addressed to: Dr. Robert Oelrichs, AIDS Molecular Biology Unit, Macfarlane Burnet Centre for Medical Research, Yarra Bend Road, Fairfield, 3078, AUSTRALIA.